EXHIBIT 1

Gene, 100 (1991) 27–38

GENE 04059

A general structure for DNA-dependent DNA polymerases

(DNA polymerase; 3′–5′ exonuclease; sequence homology; phage ϕ29; site-directed mutagenesis)

Luis Blanco, Antonio Bernad, María A. Blasco and Margarita Salas

Centro de Biología Molecular (CSIC-UAM), Universidad Autónoma, Cantoblanco, 28049 Madrid (Spain)

Received by M. Espinosa: 22 January 1991
Accepted: 24 January 1991

SUMMARY

In addition to the general 3′–5′ exonuclease domain described by Bernad et al. [Cell 59 (1989) 219–228] significant amino acid (aa) sequence similarity has been found in the C-terminal portion of 27 DNA-dependent DNA polymerases belonging to the two main superfamilies: (*i*) *Escherichia coli* DNA polymerase I (PolI)-like prokaryotic DNA polymerases, and (*ii*) DNA polymerase α-like prokaryotic and eukaryotic (viral and cellular) DNA polymerases. The six most conserved C-terminal regions, spanning approx. 340 aa, are located in the same linear arrangement and contain highly conserved motifs and critical residues involved in the polymerization function. According to the three-dimensional model of PolIk (Klenow fragment), these six conserved regions are located in the proposed polymerization domain, forming the metal and dNTP binding sites and the cleft for holding the DNA template. Site-directed mutagenesis in the ϕ29 DNA polymerase supports some of these structural predictions. Therefore, it is likely that a 'Klenow-like core', containing the DNA polymerase and 3′–5′ exonuclease activities, has evolved from a common ancestor, giving rise to the present-day prokaryotic and eukaryotic DNA polymerases.

INTRODUCTION

Structural and functional studies from several laboratories have shown a close relationship between a number of DNA-dependent DNA polymerases including both prokaryotic and eukaryotic, cellular and viral, protein-primed and RNA-primed DNA replicases (Gibbs et al., 1985; Larder et al., 1987; Knopf, 1987; Bernad et al., 1987; Wong et al., 1988), as well as the *Saccharomyces cerevisiae* REV3 DNA polymerase, probably involved in DNA repair (Morrison et al., 1989). This group, named α-like DNA polymerases, is characterized by the presence of several linearly arranged and conserved regions of aa homology, located at the C-terminal portion of the protein which were originally proposed to form the catalytic site involved in dNTP binding (Gibbs et al., 1985). In addition, the structural homology detected correlates well with similarities in sensitivity to inhibitors such as aphidicolin, BuAdATP and BuPdGTP (reviewed by Huberman, 1981; Khan et al., 1984; 1985; Blanco and Salas, 1986; Bernad et al., 1987).

Another group of DNA polymerases, also related by structural and functional similarities, is named *E. coli* PolI-like DNA polymerases. In this case, the similarities in the C-terminal portion of several enzymes from this group (Ollis et al., 1985b; López et al., 1989; Lawyer et al., 198[illegible];

Correspondence to: Dr. M. Salas, Centro de Biología Molecular (CSIC-UAM), Universidad Autónoma, Cantoblanco, 28049 Madrid (Spain) Tel. (34-1)3978435; Fax (34-1)3974799.

* Dedicated to Severo Ochoa on the occasion of his 85th birthday

Abbreviations: aa, amino acid(s); *Ac*MNPV, *Autographa californica* monocuclear polyhedrosis virus; α-human, human DNA polymerase α; α-yeast, yeast DNA polymerase α; *Bs*(III), *Bacillus subtilis* DNA polymerase III; BuAdATP, butylanilino deoxyadenosine 5′-triphosphate; BuPdGTP, butylphenyl deoxyguanosine 5′-triphosphate; dNMP, deoxynucleoside monophosphate; dNTP, deoxynucleoside triphosphate; δ-yeast, yeast DNA polymerase δ; EBV, Epstein–Barr virus; *Ec* eps, *E. coli* DNA polymerase III, epsilon subunit; HCMV, human cytomegalovirus; HSV, herpes simplex virus; nt, nucleotide(s); PAA, phosphonoacetic acid; PolI, *E. coli* DNA polymerase I; PolIk, Klenow (large) fragment of PolI; *Spn*, *Streptococcus pneumoniae*; *Taq*, *Thermus aquaticus*; VZV, varicella zoster virus; wt (or WT), wild type. Mutations are indicated by original aa (in single-letter notation), its position and the replacing aa: i.e., Y390S = Tyr^{390} → Ser.

Leavitt and Ito, 1989) support a three-dimensional structure homologous to that of PolIk (Ollis et al., 1985a).

Recently, we have shown that these two groups are interrelated by the presence of three highly conserved segments which are located in the N-terminal portion of each DNA polymerase. Based on site-directed mutagenesis studies and on the extrapolation to the x-ray structure of PolIk, we have proposed that these three segments form a highly conserved three-dimensional domain containing the 3′–5′ exonuclease active site (Bernad et al., 1989). On the other hand, it is generally accepted that the α-like and PolI-like DNA polymerases are not significantly related in the aa sequences involved in the polymerase function, located in the C-terminal portion of the polypeptide (Ollis et al., 1985b; Earl et al., 1986; Jung et al., 1987; Bernad et al., 1987; Spicer et al., 1988; Wong et al., 1988). However, as shown in this paper, structural features as the metal- and dNTP-binding sites and the regions which form the cleft for holding the DNA template are shared by enzymes belonging to both groups. The catalytic importance of several of the most conserved regions of the C-terminal portion, tested by site-directed mutagenesis in the ϕ29 DNA polymerase, support some of these structural predictions. Taking into account these structural and functional data, the modular organizations of enzymatic activities in prokaryotic and eukaryotic DNA polymerases are compared.

RESULTS AND DISCUSSION

(a) Identification of conserved aa regions from sets of sequences

Due to the difficulty of obtaining computer-derived multiple alignments when applied to proteins bearing little homology, we considered four subsets of DNA polymerases based on previous reports showing structural homology in the C-terminal portion of several groups or pairs of DNA polymerases. Group A, or PolI-like DNA polymerases, including prokaryotic enzymes as PolI, *Bs*(III), *Ec* eps, and those from Spn, Taq, and phages T5, T7 and SPO2. Group B, or eukaryotic-viral DNA polymerases, including those from HSV type 1, HCMV, EBV, VZV, *Ac*MNPV, vaccinia virus, fowlpox virus, and phage T4. This latter enzyme, although prokaryotic, has been included in this group due to the sensitivity to several inhibitors such as aphidicolin and BuPdGTP and to the presence, in its primary structure, of regions of striking similarity with animal virus DNA polymerases, as it was first described by Bernad et al. (1987). Group C, or cellular DNA polymerases, including α-human, α-yeast, δ-yeast, and REV3 from yeast. The nomenclature adopted for this group corresponds to the new one proposed by Burgers et al. (1990). Group D, or DNA polymerases from terminal protein-containing genomes, the most heterologous, including prokaryotic and eukaryotic enzymes as those from adenovirus type 2, plasmids pGKL2 and pGKL1 from yeast, plasmid pClK1 from fungi, S1 mitochondrial DNA from maize, plasmid pAI2 from fungi, phage PRD1, phage M2, and phage ϕ29. Several DNA polymerases belonging to groups B, C and D were previously classified as a α-like DNA polymerases (Bernad et al., 1987; 1989; Wong et al., 1988).

After optimal alignment of the sequences corresponding to each subset, the multiple alignment of the four subsets was carried out, proceeding orderly from the most related groups to the least (C-B-D-A). Binary alignments were carried out by computer analysis using the programs: BESTFIT, PRETTY and GAP from the UWGCG (University of Wisconsin Genetics Computer Group; Devereaux et al., 1984). This progressive alignment of the four subsets of DNA polymerases allowed us to better consider: (1) the existence and relative location of discrete segments of aa similarity, (2) the variability occurring in the conserved regions, and (3) the presence of specific conserved residues corresponding to a particular subset.

N-terminal domain. Using this progressive alignment method, the three conserved regions ExoI, ExoII and ExoIII, previously described in 19 α-like and PolI-like DNA polymerases (Bernad et al., 1989; where original references are reviewed), were detected in other recently reported sequences of DNA polymerases belonging to the four subsets (Fig. 1), as phage T5 DNA polymerase (group A; Leavitt and Ito, 1989), fowlpox virus DNA polymerase (group B; Binns et al., 1987), δ-yeast (Boulet et al., 1989) and REV3 (Morrison et al., 1989) DNA polymerases (group C), and pGKL2 (Tommasino et al., 1988), pClK1 (Oeser and Tudzinsky, 1989), pAI2 (Kempken et al., 1989) and M2 (Matsumoto et al., 1989) DNA polymerases (group D). The inclusion of the eight new sequences of DNA polymerases allowed us to improve the alignment corresponding to the ExoIII segment of T4, PRD1 and ϕ29 DNA polymerases, with respect to that previously reported (Bernad et al., 1989); in addition, the ExoIII segment of yeast polα was changed based on significant homology with human polα (T.S.-F. Wang, personal communication). Therefore, we propose that the critical residues homologous to Y497 and D501 from PolI (described in section c) are: Y320 and D324 (T4 DNA polymerase); Y145 and D149 (PRD1 DNA polymerase); Y165 and D169 (ϕ29 DNA polymerase). As shown in Fig. 1, the aa sequence of the 27 DNA polymerases compared contains regions ExoI, ExoII and ExoIII in the same linear arrangement (see also Fig. 5), supporting the idea that these three regions are forming a structurally and functionally conserved three-dimensional domain (Bernad et al., 1989). Table I shows the % similarity (averaged among the different pairs of

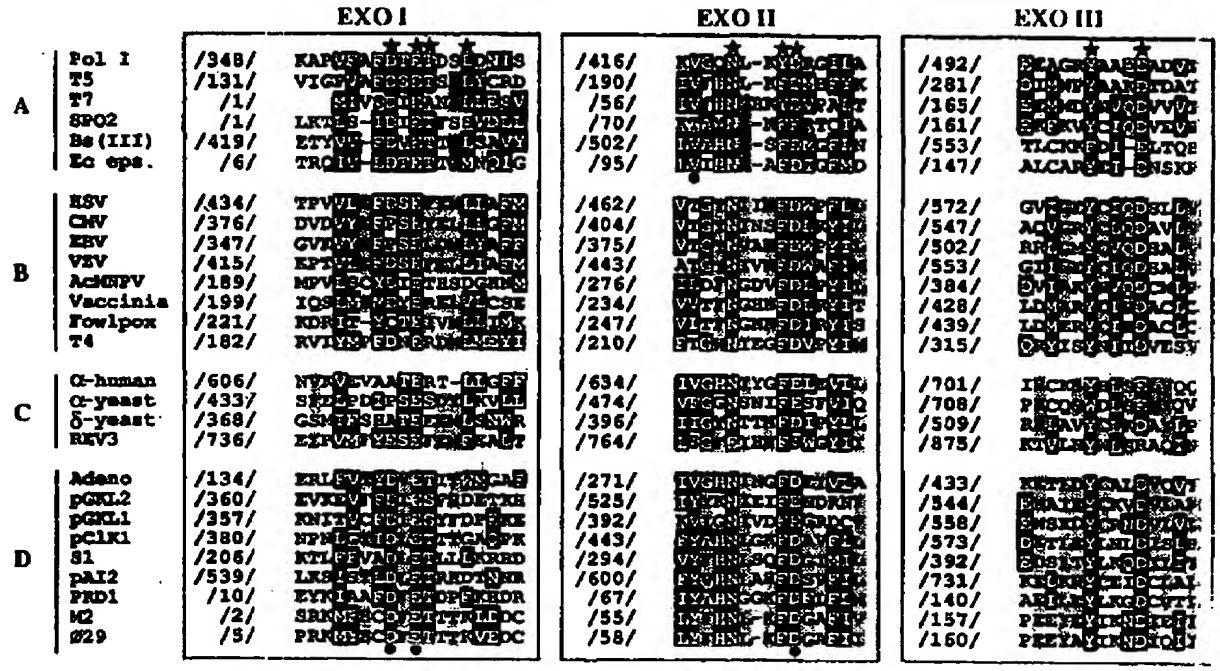

Fig. 1. Highly conserved N-terminal regions within DNA-dependent DNA polymerases. The different DNA polymerases considered were grouped as follows: A, PolI-like DNA polymerases; B, viral α-like DNA polymerases; C, cellular α-like DNA polymerases; D, protein-primed α-like DNA polymerases. DNA polymerases nomenclature is given in the text. The three highly conserved regions ExoI, ExoII and ExoIII, previously reported by Bernad et al. (1989), are indicated. Numbers between slashes indicate the aa position relative to the N-terminal end of each DNA polymerase. Relevant aa similarity among the different groups is indicated in white letters; other similarities are indicated by grey boxes. The following conservative aa were considered: S and T; A and G; K, R and H; D, E, Q and N; I, L, M, V, Y and F. Stars indicate specially conserved residues and/or PolI residues involved in exonucleolytic catalysis (see section e for details). The location of the φ29 DNA polymerase/exonuclease-deficient mutants D12A, E14A, and D66A (Bernad et al., 1989), and that of the *E. coli* DNA polymerase III (ε subunit) exonuclease-deficient mutant *dnaQ*49 (V96G) (Echols et al., 1983), are indicated with dots.

DNA polymerases) corresponding to the three Exo-regions aligned in Fig. 1, for each group of DNA polymerases.

C-terminal domain. Taking the ExoIII segment as the starting point for aligning the C-terminal domain, we found 22 segments of significant similarity for group A DNA polymerases, 21 segments for group B, 24 segments for group C, and 21 segments for group D. When the four groups were considered for multiple alignment it was possible to align 18 segments present in all enzymes belonging to the four groups. Fig. 2 shows the six most conserved regions corresponding to the multiple alignment between PolI-like enzymes (group A) and α-like DNA polymerases (groups B, C, and D). In addition, these six regions have the greatest homology in each DNA polymerase group. These

TABLE I

Amino acid similarity in the most conserved N-terminal and C-terminal regions of DNA-dependent DNA polymerases

Group*	N-terminal			C-terminal					
	I	II	III	1	2a	2b	3	4	5
A	43.3	54.3	45.2	48.7	41.6	53.0	61.5	52.0	6 .8
B	53.6	69.5	54.3	61.0	61.0	53.2	60.1	53.6	6 .7
C	40.7	57.8	29.8	74.4	63.0	52.6	66.6	62.5	4 .0
D	43.2	64.0	52.5	44.8	57.5	42.1	46.9	37.4	2 .0

* The different DNA polymerase groups (A, B, C and D), and the highly conserved regions (I, II, III, 1, 2a, 2b, 3, 4 and 5) were defined as indicated in section a and Fig. 1. Numbers indicate the % similarity averaged among the different pairs of DNA polymerases belonging to each group.

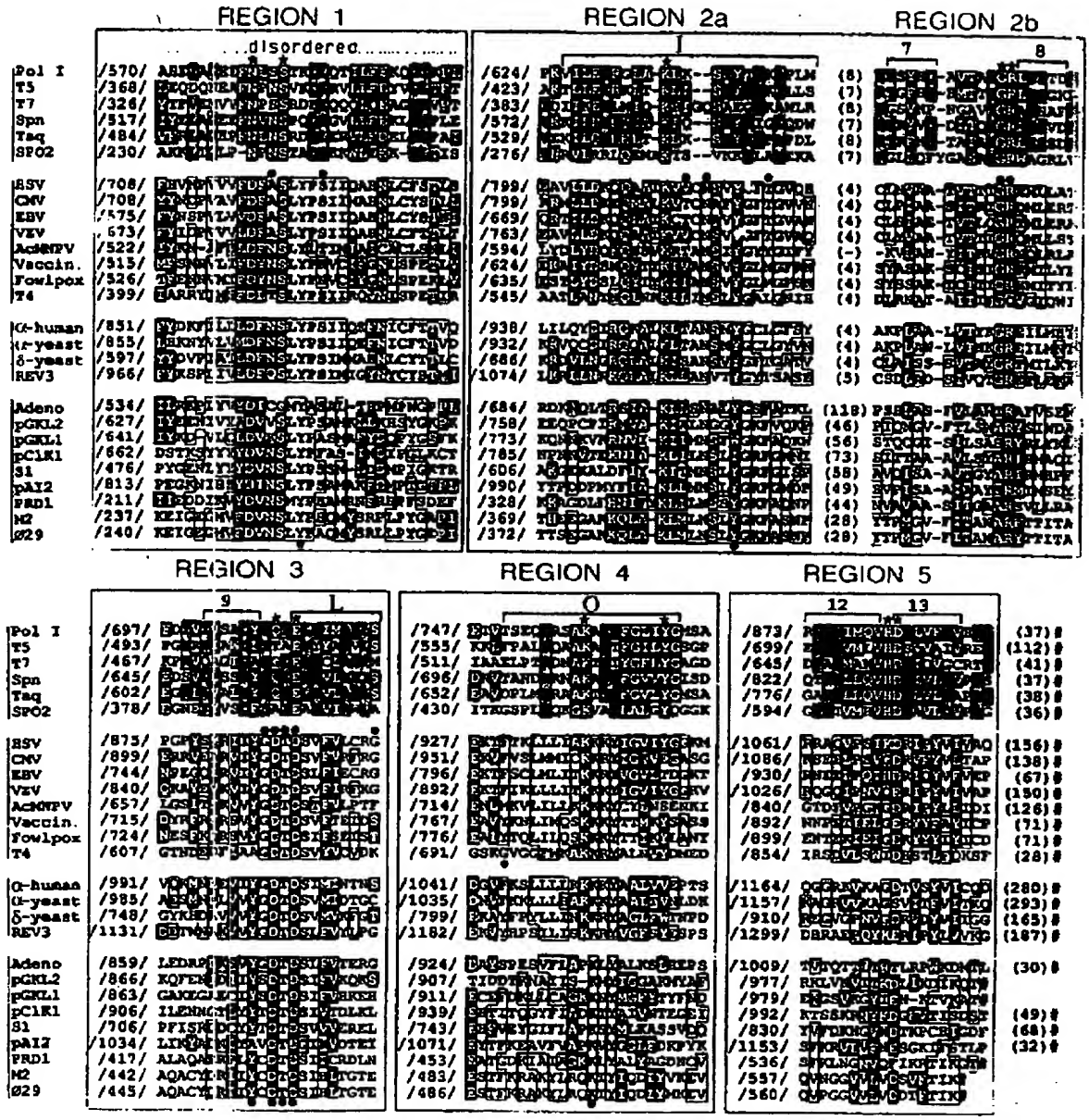

Fig. 2. Highly conserved C-terminal regions within DNA-dependent DNA polymerases. The different DNA polymerases considered were grouped as in Fig. 1. DNA polymerase nomenclature is given in section a. The six most highly conserved regions 1, 2a, 2b, 3, 4 and 5 are indicated. Numbers between slashes indicate the aa position relative to the N-terminal end of each DNA polymerase. Numbers in parentheses indicate the number of aa residues either connecting regions 2a and 2b or from the end of region 5 to the C-terminal end (#). Based on PolIk structural data (Ollis et al., 1985a), the disordered regions, α-helices (lettered) and β-strands (numbered) are indicated. Relevant aa similarity among the different groups is indicated in white letters, whereas that particularly corresponding to group A (PolI-like) and to groups B, C, and D (α-like), is indicated in grey and empty boxes, respectively. The conservative aa considered were as indicated in Fig. 1. Stars indicate specially conserved residues and/or PolI residues involved in DNA, metal and dNTP binding (see section c for details). The location of HSV DNA polymerase mutations showing altered sensitivity to PAA, acyclovir and aphidicolin: A719V, S724N, V813M, N815S, T821M, G841S and R842S (Larder et al., 1987; Knopf, 1987; Gibbs et al., 1988), and that of ϕ29 DNA polymerase mutants Y254F, Y390S and Y390F (described in section e# and Fig. 4) are indicated by dots in regions 1, 2a and 2b. The location of HSV-DNA polymerase mutations G885R, D886N, T887I, D888A and G896V (Dorsky and Crumpacker, 1990), ϕ29 DNA polymerase mutations Y454F, D456G, T457P and D458G (Bernad et al., 1989), K498T and K498R (unpublished results), and that of T4 DNA polymerase mutator mutant tsL88 (G694S) (Hershfield, 1973), are also indicated by dots in regions 3 and 4.

regions were numbered as they occurred from the N terminus to the C terminus, however, the presence of a specific insertion in region 2 of protein-primed DNA polymerases (group D) led us to divide region 2 into regions 2a and 2b.

Region 1, spanning 30 aa, contains in groups B, C and D the highly conserved motif 'D—SLYP' previously described. The multiple alignment shows that the Ser residue corresponding to this motif (marked by a star in Fig. 2) is also highly conserved in PolI-like DNA polymerases (group A). Furthermore, the highly conserved aspartate in α-like DNA polymerases is substituted by a highly conserved asparagine (marked by a star in Fig. 2) in PolI-like DNA polymerases. **Region 2a**, spanning 27 aa, contains in groups B, C and D the highly conserved motif 'K—NS-YG' previously described. The multiple alignment shows that the Lys residue corresponding to this motif (marked by a star in Fig. 2) is also highly conserved in PolI-like DNA polymerases, although in this case the use of small gaps was necessary for optimal alignment. Between regions 1 and 2a, an insertion of about 50 aa, forming a highly conserved region, is only present in protein-primed DNA polymerases (unpublished results). **Region 2b**, spanning 19 aa residues, contains in all four groups a highly conserved Arg preceded by an invariant Gly in groups A, B and C (marked by stars in Fig. 2); this Gly is substituted by a relatively conserved Ala in group D. Regions 2a and 2b are connected by a short number of aa residues in groups A, B, and C, whereas in group D, a larger number of aa are forming a specific conserved region, previously reported by Bernad et al. (1987). **Region 3**, spanning 22 aa, contains in groups B, C and D the highly conserved motif 'YGDTDS', although the Gly residue corresponding to this motif strongly varies in group D DNA polymerases. The multiple alignment shows that the second Asp of this motif aligns with an invariant Glu in group A (marked by a star in Fig. 2), whereas the first Asp of the 'YGDTDS' motif corresponds to a Gln residue in PolI (marked by a star in Fig. 2), *Spn* and *Taq* DNA polymerases. **Region 4**, spanning 24 aa, contains in group A an invariant Arg (boxed in Fig. 2), and two invariant Lys and Tyr residues (indicated by stars in Fig. 2). The multiple alignment shows that groups B, C and D share an invariant Lys and Tyr residue (boxed in Fig. 2), in addition to another highly conserved Lys residue (groups B and C; printed in white letters in Fig. 2) that aligns with that corresponding to group A. It is worth noting the high similarity in this region between PolI and both HSV and VZV DNA polymerases. **Region 5**, spanning 19 aa, contains in group A the invariant motif 'VHD'. The multiple alignment shows that groups B, C and D contain a highly conserved Asp or Glu in the position corresponding to the invariant Asp in group A (indicated by a star in Fig. 2). It is worth noting the high similarity, in the central portion of this region, between the EBV DNA polymerase and the PolI-like DNA polymerases.

The fact that the protein alignment corresponding to these evolutionary distant 27 DNA polymerases is colinear (see also Fig. 5), supports the idea that these six regions are forming a structurally and functionally conserved three-dimensional domain. Table I shows the % similarity (averaged among the different pairs of DNA polymerases) corresponding to the six C-terminal regions aligned in Fig. 2, for each group of DNA polymerases.

(b) Three-dimensional structure prediction

PolIk, considered as the prototype of multifunctional enzyme involved in the repair and replication of DNA (Kornberg, 1980) is, to date, the only DNA polymerase whose three-dimensional structure is known (Ollis et al., 1985a). The crystal analysis of PolIk shows that the polypeptide is folded into two distinct structural domains of approx. 200 and 400 aa (Fig. 3). The smaller domain (N-terminal) contains the 3′–5′ exonuclease activity, whereas the larger domain (C-terminal) contains the polymerization active site. The larger domain contains a cleft, about 20–24 Å wide and 25–35 Å deep, with the appropriate dimensions for holding double-stranded B DNA (reviewed by Joyce and Steitz, 1987). By extrapolation to the x-ray structure of PolIk, we have recently proposed that the 3′–5′ exonuclease domain is structurally and functionally conserved among prokaryotic and eukaryotic DNA polymerases; in particular, the three highly conserved regions ExoI, ExoII and ExoIII (light-shaded areas in Fig. 3), forming the 3′–5′ exonuclease active site (Bernad et al., 1989).

On the basis of the homology detected among the C-terminal portion of PolI and other DNA polymerases belonging to groups A, B, C and D, the location of the six most conserved regions (aligned in Fig. 2) was analyzed by extrapolation to the three-dimensional structure of PolIk. As shown in Fig. 3, region 1 lies in a portion not sufficiently ordered in the crystal (tentatively depicted with dashed lines in Fig. 3), which has been proposed to close off the fourth side of the cleft (Ollis et al., 1985a). Regions 2a and 4, corresponding to α-helices I and O, respectively, form two opposite walls of the cleft, whereas regions 3 (β-strand 9 followed by α-helix L), 2b (β-strands 7 and 8), and 5 (β-strands 12 and 13) correspond to the five-stranded antiparallel β-sheet forming the floor of the cleft. The other twelve regions of lower sequence similarity also correspond to three-dimensional portions having a defined secondary structure; in general, the variable regions among the DNA polymerases correspond to random-coiled portions connecting well-ordered regions (unpublished results). Exceptionally, some turns are specially conserved, as those existing between β-strand 9 and α-helix L (region 3) and between β-strands 7 and 8 (region 2b) or β-strands 12 and 13 (region 5). In summary, the fact that the six most conserved regions of aa sequence similarity are concentrated around

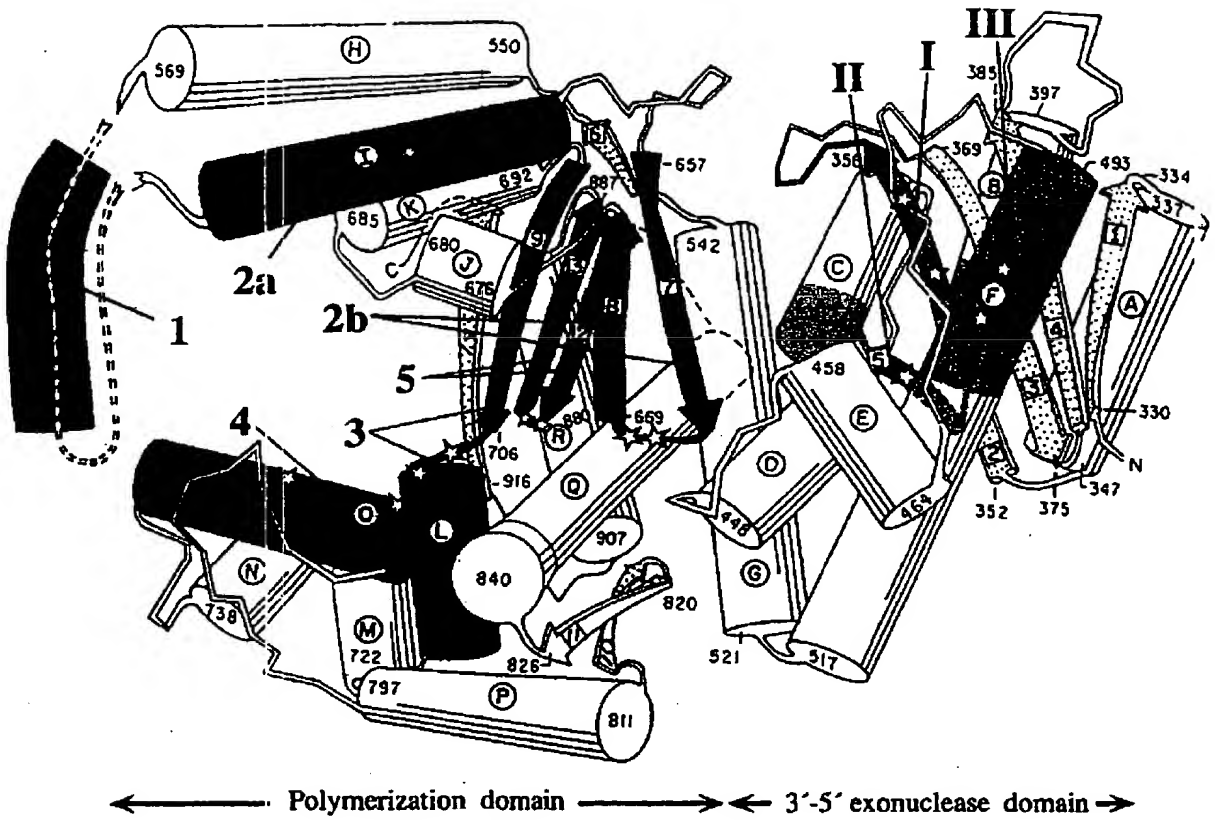

Fig. 3. A Klenow-like general structure for DNA-dependent DNA polymerases. The representation of the three-dimensional structure of PolIk is from Ollis et al. (1985a), with slight modifications. Regions that form α-helices are represented by cylinders (lettered), and those forming β-sheets by stippled arrows (numbered); the disordered region spanning aa 569–626 is indicated by dashed lines. The N-terminal portion (residues 330–520) of the crystal structure contains the 3′–5′ exonuclease, particularly involving the three highly conserved regions ExoI, ExoII and ExoIII (I, II and III; indicated with light-shaded areas aligned in Fig 1). The C-terminal portion (aa 521–928), containing the polymerization activity, involves the six highly conserved C-terminal regions 1, 2a, 2b, 3, 4 and 5 (dark-shaded areas shown in Fig. 2). Stars mark the location of aa residues involved in the 3′–5′ exonuclease activity (metal binding and catalysis) and DNA polymerization activity (DNA, metal and dNTP binding and catalysis) of PolI. These positions correspond to highly conserved aa marked by stars in Figs. 1 and 2.

the cleft that is proposed to bind DNA, emphasizes its functional importance and the conservation of this structural feature throughout the evolution of this key enzyme for the process of DNA replication.

(c) Functional significance of the conserved regions

The structural similarity detected among PolI and other prokaryotic and eukaryotic DNA polymerases (Bernad et al., 1989; this paper) allows us to generalize some aspects of the DNA polymerase function which are suggested from the unique shape of the tertiary structure of the PolIk. Another important line of evidence comes from the use of specific reagents or modified substrates, as well as from site-directed mutagenesis studies of highly conserved or putatively important aa residues in some of the related DNA polymerases.

(*i*) *N-terminal domain.* A structurally independent domain, as that of PolIk, contains the 3′–5′ exonuclease active site in other prokaryotic and eukaryotic DNA polymerases. This evolutionarily conserved active site is mainly formed by the highly conserved regions ExoI, ExoII and ExoIII (Bernad et al., 1989; Fig. 1). In agreement with this hypothesis, PolI, T7 and ϕ29 DNA polymerase mutant proteins containing aa changes in the exonuclease active site, made by site-directed mutagenesis, are devoid of exonuclease activity but retain full polymerase activity (Derbyshire et al., 1988; Tabor and Richardson, 1989; Bernad et al., 1989). The *dnaQ*49 mutation, located at position 96 in the ExoII region of the ε subunit of the *E. coli* DNA PolIII holoenzyme (indicated by a dot in Fig. 1), produces a strong mutator phenotype (Horiuchi et al., 1978) and a defective 3′–5′ exonuclease activity (Echols

et al., 1983). In addition, the N-terminal clustering of T4 DNA polymerase mutator mutants (Reha-Krantz, 1988; 1989) was also informative to localize the exonuclease domain (Bernad et al., 1989). The more extensive homology presented in this paper strengthens the importance of the Asp and Glu (involved in metal binding in PolI; Derbyshire et al., 1988) of region ExoI (indicated with stars in Figs. 1 and 3) in groups A and D, whereas only the Glu is invariant in groups B and C. Thr^{358} of PolI (indicated by a star in Fig. 1), which buries the 3'-OH group of the DNA substrate by formation of an H-bond through its backbone amide (Freemont et al., 1988) is also highly conserved in groups A and D, whereas Leu^{361} of PolI (indicated by a star in Fig. 1), whose side chain interacts with the base moiety of the last nt of the DNA substrate (Freemont et al., 1988), is conserved in groups A, B and C. In region ExoII, the Asp^{424} of PolI (indicated by a star in Figs. 1 and 3), which is involved in metal binding (Derbyshire et al., 1988), is always preceded by an aromatic residue (Y or F; indicated by a star in Fig. 1), and is highly conserved in all four groups, although it can be substituted by a Glu residue. Interestingly, with the exception of REV3 DNA polymerase, all the enzymes belonging to the four groups contain, in the ExoII region, an invariant Asn residue (indicated by a star in Fig. 1) whose particular role in the exonuclease active site is presently unknown. In region ExoIII, the Tyr^{497} of PolI (indicated by a star in Figs. 1 and 3), proposed to be involved in exonucleolytic catalysis (Freemont et al., 1988), is almost invariant in the four groups, and the Asp^{501} of PolI (indicated by a star in Figs. 1 and 3), involved in metal binding (Derbyshire et al., 1988), is highly conserved in groups A, B and D. These data do not support the aa sequence alignment reported by Matsumoto et al. (1989).

(*ii*) *C-terminal domain.* A structurally independent domain, as that of the PolIk, is proposed to contain the polymerization active site in prokaryotic and eukaryotic DNA polymerases. An evolutionarily conserved 'DNA polymerase core' would be formed by the highly conserved regions 1, 2a, 2b, 3, 4 and 5 (Fig. 2).

(*iii*) *DNA binding.* Region 1 is located in a subdomain of PolIk that is disordered in the electron-density map (Fig. 3). It has been suggested that this flexible subdomain may bind to the DNA, thereby closing off the cleft and allowing the protein to surround the bound DNA (Ollis et al., 1985a). In fact, this region corresponds to the most or one of the most flexible portions of the different DNA polymerases compared in Fig. 2 (unpublished results). It has been proposed that the function of this flexible subdomain, containing region 1, may be to increase the processivity of the enzyme by slowing its rate of dissociation from the DNA relative to translocation and further nt addition (Ollis et al., 1985a). Region 2a, located immediately after the disordered domain, is structured as a long α-helix (I-helix in PolIk; Fig. 3) which forms one side of the cleft. The aa sequence of PolI contains, in this region, Lys^{635} (marked by a star in Figs. 2 and 3) which has been directly involved in DNA binding and processivity of the polymerization reaction (Basu et al., 1988). Interestingly, a corresponding Lys residue is invariant in both PolI-like and α-like DNA polymerases (Fig. 2). In the latter, this Lys forms part of the highly conserved motif 'K—NS-YG' previously described. Therefore, these structural and functional data suggest that regions 1 and 2a are probably involved in DNA binding and processivity, the coupling of synthesis to movement along the template.

To test the above hypothesis, site-directed mutagenesis in the highly conserved Tyr residues in regions 1 and 2a of the ϕ29 DNA polymerase was carried out. These Tyr residues were selected for mutagenesis on the basis of the apparent similarity of the motifs enclosed in these regions: 'D-NSLYP' (region 1) and 'K—NS(L/V)YG' (region 2a). Tyr^{254} (region 1) was changed into Phe, whereas Tyr^{390} (region 2a) was changed into Ser or Phe, and the mutant proteins were overproduced, purified and assayed for 3'–5' exonuclease, nonprocessive and processive DNA polymerization, and protein-primed initiation. As shown in Fig. 4B, the Y254F and Y390F mutations strongly affected the processive elongation of the ϕ29 DNA polymerase (Blanco et al., 1989), whereas the Y390S mutation essentially had no effect. Fig. 4A shows that nonprocessive elongation was only affected by the Y390F mutation; however, when the turnover of dNTPs to dNMPs coupled to this nonprocessive assay was analyzed, an abnormally high 3'–5' exonuclease activity was detected for the three mutants, Y254F, Y390F and Y390S, being 62-fold, 16-fold and 22-fold higher, respectively, than that of the wt protein (Fig. 4C). The fact that the 3'–5' exonuclease activity, assayed in the absence of polymerization, was normal for the three mutants (not shown) indicates that the increased turnover observed at the DNA 3'-terminus is probably reflecting a slow rate of translocation of the mutant ϕ29 DNA polymerases along the DNA template. Therefore, and in agreement with the proposed structural model, regions 1 and 2a are probably defining DNA polymerase regions that interact with the DNA. These results also confirm and limit the proposed location of the exonuclease domain, supporting its physical separation from the polymerase domain (Bernad et al., 1989).

Interestingly, the Y254F mutation (region 1) strongly affected the protein-primed initiation reaction by decreasing about 30-fold the formation of the terminal protein (p3)-dAMP initiation complex (Fig. 4D). This result, together with the neighbouring location of a specific region of about 50 aa residues preceding region 2a, only present in protein-primed DNA polymerases (unpublished results),

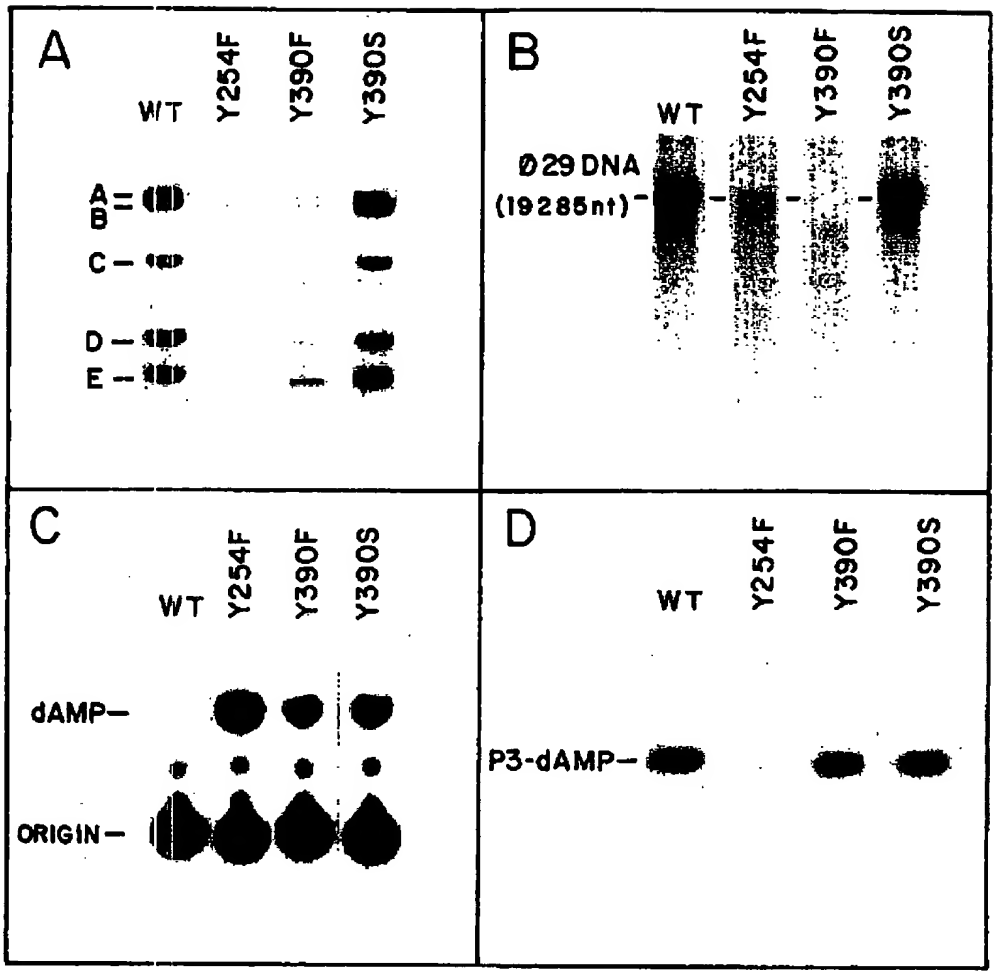

Fig. 4. Effect of site-directed mutations in regions 1 and 2a of the ϕ29 DNA polymerase. The ϕ29 DNA polymerase mutants Y254F (region 1), Y390S and Y390F (region 2a) were obtained, expressed and purified as described by Bernad et al. (1990a). (A) Nonprocessive DNA polymerization assay. The filling-in reaction on *Eco*RI-digested ϕ29 DNA (0.5 μg), assayed in the presence of either wt or mutant ϕ29 DNA polymerase (1 ng), was carried out essentially as described by Bernad et al. (1990b); the activity values, indicated as % of that of the wt protein, were: Y254F (80%); Y390F (20%); Y390S (100%). (B) Processive DNA polymerization assay. The replication of ϕ29 DNA-protein p3 (0.5 μg), assayed in the presence of either wt or mutant ϕ29 DNA polymerase (1 ng), was carried out essentially as described by Bernad et al. (1990b); the activity values, expressed as % of that of the wt protein were: Y254F (5%), Y390F (<2%), Y390S (60%). (C) Turnover coupled to the nonprocessive DNA polymerization assay. The dAMP turnover was determined by polyethyleneimine-cellulose thin-layer chromatography and further autoradiography of samples taken immediately after the DNA polymerization reaction. The chromatogram was developed with 0.15 M Li·formate pH 3.0, conditions in which the 5'-dAMP migrates, whereas the DNA and the unincorporated dNTPs remain at the origin. Turnover was quantitated by counting the Cerenkov radiation corresponding to the 5'-dAMP spot; the activity values, relative to that of the wt protein were: Y254F (62-fold); Y390F (16-fold); Y390S (22-fold). (D) Protein-primed initiation assay. The formation of the p3-dAMP covalent complex, assayed in the presence of ϕ29 DNA-protein p3 (0.5 μg) as template, and either wt or mutant ϕ29 DNA polymerase (1 ng), was carried out essentially as described by Bernad et al. (1990b), and analyzed by 0.1% SDS–10% polyacrylamide gel electrophoresis and autoradiography; the activity values, expressed as % of that of the wt protein, were: Y254F (3%); Y390F (75%); Y390S (89%).

suggests that region 1 is involved, in addition to DNA-binding, in interaction with the priming protein. Taking into account the proposed flexibility of the structure in the neighbourhood of region 1, it is tempting to speculate that the strong interaction detected between the ϕ29 DNA polymerase and terminal protein (Blanco et al., 1987) is achieved by location of the latter inside the cleft and further stabilized by interaction with a slightly 'adapted' flexible domain.

(*iv*) *Metal binding.* Recent crystallographic studies of PolIk with dCTP have shown that PolI residues Gln^{708} and Glu^{710} (region 3) are involved in the interaction with an

associated Mg^{2+} ion (L. Beese, J. Friedman and T. Steitz, personal communication). Interestingly, the corresponding region of α-like DNA polymerases contains the highly conserved 'YGDTDS' motif, proposed to be involved in metal binding and catalysis (Argos 1988; Bernad et al., 1990b). Site-directed mutagenesis in the ϕ29 DNA polymerase indicated that the Thr and second Asp of the motif seem to be the most critical residues for both the protein-primed initiation and polymerization function of the ϕ29 DNA polymerase (Bernad et al., 1990b). In agreement with their relative functional importance, the first Asp of the 'YGDTDS' motif corresponds to the relatively conserved Gln^{708} of PolI (indicated by a star in Figs. 2 and 3), whereas the critical second Asp corresponds to Glu^{710} of PolI (indicated by a star in Figs. 2 and 3), invariably conserved in group A DNA polymerases. The involvement of region 3 of PolI in metal binding does not support the alignment reported by Delarue et al. (1990), in which no Gln, Glu or Asp residues would be present in the corresponding positions of α-like DNA polymerases. Furthermore, the change of ϕ29 DNA polymerase residue Cys^{455} into the consensus Gly of the motif, produced an abnormal behaviour in the usage of activating metal ions by the ϕ29 DNA polymerase (Bernad et al., 1990b). The functional importance of region 3 for polymerase activity has been also demonstrated in the HSV DNA polymerase; in this case, single changes in the 'GDTD' sequence destroyed the polymerization activity (Dorsky and Crumpacker, 1990). Therefore, all these data strongly suggest that the aspartates contained in the 'YGDTDS' motif play a direct role in metal binding in DNA polymerases belonging to groups B, C and D.

In PolIk, region 5 corresponds with two antiparallel β-strands (12 and 13) connected by a turn. One of the residues located in this turn is Asp^{882} (indicated by a star in Figs. 2 and 3), shown to be directly interacting with the polymerization metal ion, in addition to Gln^{708} and Glu^{710} residues described before (L. Beese, J. Friedman and T.A. Steitz, personal communication). The catalytic significance of this residue is consistent with the high conservation of this Asp (or Glu) in most of the DNA polymerases compared (Fig. 2). In the case of group D DNA polymerases, the conservation of region 5 is significantly lower in comparison with the other groups (Table I). Therefore, the higher divergence of this region in group D that could be related to its proximity to the very C terminus, leads to take with caution, in the case of the protein-primed DNA polymerases, the functional significance inferred for this region.

(*v*) *dNTP binding.* Joyce and Steitz (1987) have suggested that the dNTP-binding site of PolIk lies between the C-terminal end of the O-helix, the N-terminal end of the Q-helix and the bed formed by strands 7, 8, 12 and 13 (Fig. 3). This location is consistent with photoaffinity labelling studies showing a direct role of PolI Tyr^{766} (Rush and Konigsberg, 1990), located at the C-terminal end of O-helix, and PolI His^{881} (Pandey et al., 1987), which is located in the turn between β-strands 12 and 13. Additionally, pyridoxal phosphate, which binds competitively to the dNTP site, reacts with PolI Lys^{758}, also in the O-helix (Basu and Modak, 1987). According to the alignment shown in Fig. 2, region 4 corresponds to the O-helix in PolI. In agreement with their functional importance, the residues corresponding to PolI Lys^{758} and Tyr^{766} (indicated by stars in Figs. 2 and 3) are invariant in the DNA polymerases belonging to group A. In the case of groups B, C and D, the corresponding region also contains highly conserved Lys and Tyr residues (boxed in Fig. 2) which could be the functional counterparts of Lys^{758} and Tyr^{766} of PolI. In agreement with that, substitution of the corresponding Lys^{498} residue in region 4 of the ϕ29 DNA polymerase into Thr or Arg (Fig. 2) completely destroyed the polymerization activity of the enzyme (unpublished results). Secondary structure analysis predicts, in all the DNA polymerases compared including PolI, that region 4 is structured as an α-helix at its N-terminal two thirds, whereas the C-terminal third has a clear hydrophobic character (not shown). These data suggest that a hydrophobic environment is probably important to favour dNTP binding. In region 5, and also in agreement with its functional importance, an His residue corresponding to PolI His^{881} (indicated by a star in Figs. 2 and 3) is invariably present in group A DNA polymerases; however, with the exception of the EBV DNA polymerase, this residue is not conserved in groups B, C and D. According to the three-dimensional model, the spatial proximity of the residues involved in both metal and dNTP binding (Fig. 3), agrees with the functional requirement of the dNTP substrates in their metal-quelated form. This intimal relationship is also reflected by the conservation of the distance between regions 3, 4 and 5 in most of the evolutionarily distant DNA polymerases compared (Figs. 2 and 5).

Recently, a 50-aa peptide encompassing the PolI O-helix was shown to bind both nt substrates and duplex DNA (Mullen et al., 1989). This fact, and the structural location of O-helix (forming one side of the cleft), leads us to speculate that region 4 could have a dual role in both dNTP and DNA binding, being important to adjust the incoming nt at the proper distance with respect to the primer-terminus and template-nt. Thus, the coordinated binding of both substrates, DNA and dNTPs, could provide an important contribution to the insertion fidelity. In agreement with this idea, it is worth noting that the phenotype of one of the classical T4 mutator mutants (tsL88), having a reduced specificity in inserting correct nt (Hershfield, 1973), is due to the single change $Gly^{594} \rightarrow Ser$; as shown in Fig. 2 this position lies in region 4 of T4 DNA polymerase.

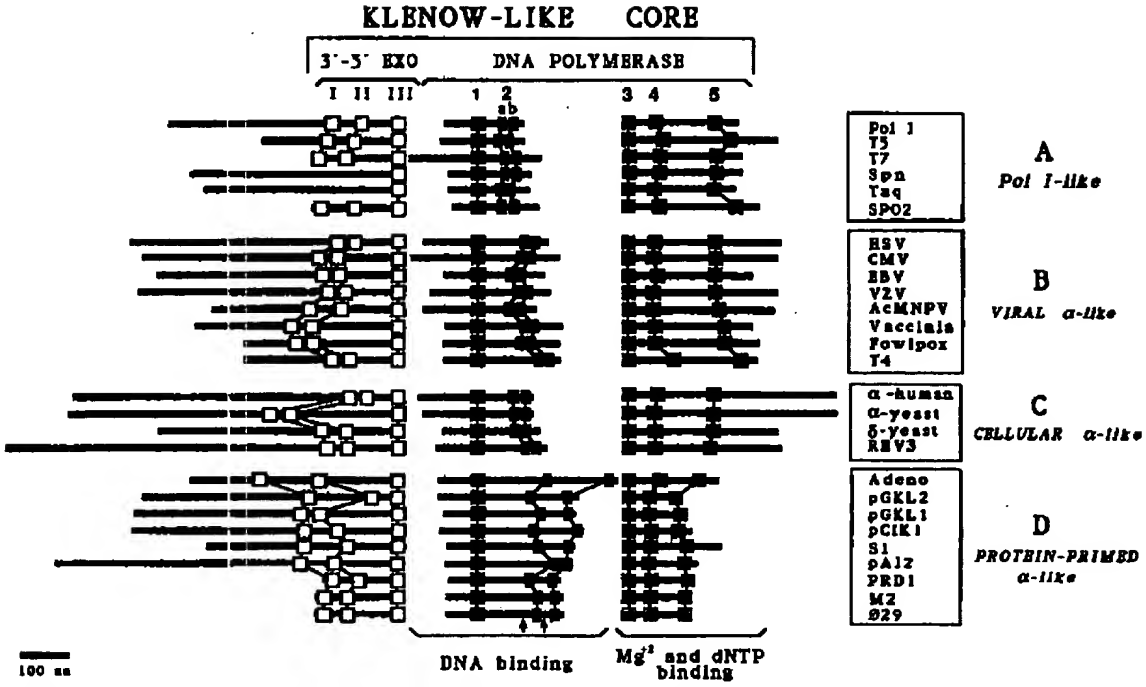

Fig. 5. Modular organization of enzymatic activities in DNA-dependent DNA polymerases. Each DNA polymerase is represented with the N terminus at the lefthand side. Open squares indicate the three highly conserved N-terminal regions ExoI, ExoII and ExoIII (shown in Fig. 1), proposed to form a general 3'–5' exonuclease-active site (Bernad et al., 1989). Blackened squares indicate the six highly conserved C-terminal regions shown in Fig. 2. Regions 1, 2a and 2b are proposed to be mainly involved in DNA binding, whereas regions 3, 4 and 5 are proposed to be mainly involved in metal and dNTP binding. The ExoIII segment was considered as an alignment axis for the N-terminal portion of each DNA polymerase. To reflect the variation in length between regions 1 and 2, and the conservation of the distance between regions 3, 4 and 5, regions 1 and 3 were considered as two independent alignment axes for the C-terminal portion of the DNA polymerases compared. Specific insertions in protein-primed DNA polymerases are indicated by arrows in the ϕ29 DNA polymerase. DNA polymerase nomenclature is given in section a. A scale bar is given at the lower left.

(d) Modular organization in DNA-dependent DNA polymerases

As shown in Fig. 5, the presence of the linearly arranged conserved regions ExoI, ExoII and ExoIII (3'–5' exonuclease), and regions 1, 2a, 2b, 3, 4 and 5 (DNA polymerization), in most of the DNA polymerases belonging to any of the four groups, suggests that these two domains correspond to an evolutionarily conserved Klenow-like DNA polymerase 'core'. As shown in Fig. 5, *Spn* and *Taq* DNA polymerases do not contain the ExoI and ExoII regions, in agreement with the fact that they have very low, if any, 3'–5' exonuclease activity (López et al., 1989; Lawyer et al., 1989). Other DNA polymerases contain N-terminal and C-terminal segments out of the proposed 3'–5' exonuclease and polymerase domains. In the case of PolI, the N terminus contains a 5'–3' exonuclease activity which can be structurally and functionally separated from the PolIk by mild proteolysis (Lehman and Chien, 1973; Jacobsen et al., 1974; Joyce et al., 1985). In *Spn* and and *Taq* DNA polymerases, the first 274 aa from the N-terminus have evident homology (36% and 37% identical residues, respectively) with the 5'–3' exonuclease domain of PolI (López et al., 1989; Lawyer et al., 1989). In agreement with these structural data, the purified pneumococcal enzyme exhibits 5'–3' exonuclease activity (López et al., 1989). Recently, an RNase H activity, which may have a similar role to that of a 5'–3' exonuclease for removing RNA primers, has been reported to be an integral part of the HSV DNA polymerase (Crute and Lehman, 1989). Based on aa sequence comparisons, this activity has been proposed to reside in the N terminus of HSV and also in VZV, EBV and HCMV DNA polymerases. In agreement with that, it has been reported that this RNase H was not only found to be associated to the 136-kDa HSV DNA polymerase, but also to a 30-kDa polypeptide that may be a proteolytic fragment analogous to the 'small' fragment (containing the 5'–3' exonuclease) of PolI (Crute and Lehman, 1989). The size of that polypeptide fits in the N-terminal region of HSV DNA polymerase located upstream from the 3'–5' exonuclease domain (Fig. 5). In the case of T7 and T5 DNA polymerases and supporting the existence of structural and functional 'modules' for the different catalytic activities of these enzymes, a separate 5'–3' exonuclease has been reported to be homologous to that associated to PolI (López

et al., 1989; Leavitt and Ito, 1989). Whether the most N-terminal and C-terminal portions of other DNA polymerases contain other enzymatic activities or have some other specific functions, such as interaction with accessory proteins, remains to be determined.

(e) Conclusions

The structural and functional analyses reported in this paper allow us to propose that the different DNA polymerases compared share a similar overall structure, as that of the PolIk; nevertheless, significant differences as the presence of specific regions or aa motifs in the proposed DNA binding, metal-binding and dNTP-binding sites, probably confer their individual characteristics of processivity and insertion fidelity.

Interestingly, in the case of the ϕ29 DNA polymerase, one of the smallest DNA polymerases having 3′–5′ exonuclease activity, RNA primers are not required for a priming protein-initiated mechanism of continuous DNA replication and this enzyme is, by itself, highly processive and able to produce strand displacement in the absence of any accessory protein (Blanco et al., 1989). The properties are likely related with the fact that its size is restricted to the proposed 'Klenow-like core'. On the other hand, the structural comparison of protein-primed DNA polymerases (group D) with those belonging to the other three groups indicates that the specific protein-priming function does not correspond to any additional DNA polymerase structural 'module'. On the contrary, the presence of specific insertions flanking region 2a (indicated with arrows in Fig. 5) suggests that the ability to use a protein as a primer has been acquired by modification of the DNA polymerization 'core' of a primordial DNA polymerase.

Awaiting the resolution of the three-dimensional structure of new DNA polymerases, the aa sequence similarities reported in this paper lead to testable predictions of critical residues for both structure and function in DNA-dependent DNA polymerases.

ACKNOWLEDGEMENTS

We thank E. Parés for the ϕ29 DNA polymerase mutants Y254F, Y390F and Y390S, and J.M. Lázaro for excellent technical assistance in ϕ29 DNA polymerase purification. We are grateful to Dr. W.F. Studier for the *E. coli* BL21(DE3) expression strain, and to Drs. S. Tabor and C.C. Richardson for plasmid pT7-3. This investigation has been aided by research Grant 5 R01 GM27242-11 from the National Institutes of Health, by Grant No. PB87-0323 from Dirección General de Investigación Científica y Técnica, and by a grant from Fundación Ramón Areces. M.A.B. was a predoctoral fellow from Dirección General de Investigación Científica y Técnica and A.B. was a postdoctoral fellow from Consejo Superior de Investigaciones Científicas.

REFERENCES

Argos, P.: A sequence motif in many polymerases. Nucleic Acids Res. 16 (1988) 9909–9916.

Basu, A. and Modak, M.J.: Identification and amino acid sequence of the deoxynucleoside triphosphate binding site in *Escherichia coli* DNA polymerase I. Biochemistry 26 (1987) 1704–1709.

Basu, S., Basu, A. and Modak, M.J.: Pyridoxal 5′-phosphate mediated inactivation of *Escherichia coli* DNA polymerase I: identification of lysine-635 as an essential residue for the processive mode of DNA synthesis. Biochemistry 27 (1988) 6710–6716.

Bernad, A., Zaballos, A., Salas, M. and Blanco, L.: Structural and functional relationships between prokaryotic and eukaryotic DNA polymerases. EMBO J. 6 (1987) 4219–4225.

Bernad, A., Blanco, L., Lázaro, J.M., Martin, G. and Salas, M.: A conserved 3′–5′ exonuclease active site in prokaryotic and eukaryotic DNA polymerases. Cell 59 (1989) 219–228.

Bernad, A., Blanco, L. and Salas, M.: Site-directed mutagenesis in the YCDTDS motif of the ϕ29 DNA polymerase. Gene 94 (1990a) 45–51.

Bernad, A., Lázaro, J.M., Salas, M. and Blanco, L.: The highly conserved amino acid sequence motif Tyr-Asp-Thr-Asp-Ser in α-like DNA polymerases is required by phage ϕ29 DNA polymerase for protein-primed initiation and polymerization. Proc. Natl. Acad. Sci. USA 87 (1990b) 4610–4614.

Binns, M.M., Stenzler, L., Tomley, F.M., Campbell, J. and Boursnell, M.E.G.: Identification by a random sequence strategy of the fowlpox DNA polymerase gene, its nucleotide sequence and comparison with other viral DNA polymerases. Nucleic Acids Res. 15 (1987) 6563–6573.

Blanco, L. and Salas, M.: Effect of aphidicolin and nucleotide analogs on the phage ϕ29 DNA polymerase. Virology 153 (1986) 179–187.

Blanco, L., Prieto, I., Gutiérrez, J., Bernad, A., Lázaro, J.M., Hermoso, J.M. and Salas, M.: Effect of NH_4^+ ions on ϕ29 DNA-protein p3 replication: formation of a complex between the terminal protein and the DNA polymerase. J. Virol. 61 (1987) 3983–3991.

Blanco, L., Bernad, A., Lázaro, J.M., Martin, G., Garmendia, C. and Salas, M.: Highly efficient DNA synthesis by the phage ϕ29 DNA polymerase. Symmetrical mode of DNA replication. J. Biol. Chem. 264 (1989) 8935–8940.

Boulet, A., Simon, M., Faye, G., Bauer, G.A. and Burgers, M.J.: Structure and function of the *Saccharomyces cerevisiae CDC2* gene encoding the large subunit of DNA polymerase III. EMBO J. 8 (1989) 1849–1854.

Burgers, P.M.J., Bambara, R.A., Campbell, J.L., Chang, L.M.S., Downey, K.M., Hübscher, U., Lee, M.Y.W.T., Linn, S.M., So, A.G. and Spadari, S.: Revised nomenclature for eukaryotic DNA polymerases. Eur. J. Biochem. 170 (1990) 1–2.

Crute, J.J. and Lehman, I.R.: Herpes simplex-1 DNA polymerase: identification of an intrinsic 5′–3′ exonuclease with ribonuclease H activity. J. Biol. Chem. 264 (1989) 19266–19270.

Delarue, M., Poch, O., Tordo, N., Moras, D. and Argos, P.: An attempt to unify the structure of polymerases. Prot. Eng. 6 (1990) 461–467.

Derbyshire, V., Freemont, P.S., Sanderson, M.R., Beese, L., Friedman, J.M., Joyce, C.M. and Steitz, T.A.: Genetic and crystallographic studies of the 3′–5′ exonucleolytic site of DNA polymerase I. Science 240 (1988) 199–201.

Devereaux, J., Haeberli, P. and Smithies, O.: A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12 (1984) 387–395.

Dorsky, D.I. and Crumpacker, C.S.: Site-specific mutagenesis of a highly conserved region of the herpes simplex virus type 1 DNA polymerase gene. J. Virol. 64 (1990) 1394–1397.

Earl, P.L., Jones, E.V. and Moss, B.: Homology between DNA polymerases of poxviruses, herpesviruses and adenoviruses: nucleotide sequence of the vaccinia virus DNA polymerase gene. Proc. Natl. Acad. Sci. USA 83 (1986) 3659–3663.

Echols, H., Lu, C. and Burgers, P.M.J.: Mutator strains of *Escherichia coli*, *mutD* and *dnaQ*, with defective exonucleolytic editing by DNA polymerase III holoenzyme. Proc. Natl. Acad. Sci. USA 80 (1983) 2189–2192.

Freemont, P.S., Friedman, J.M., Beese, L.S., Sanderson, M.R. and Steitz, T.A.: Cocrystal structure of an editing complex of Klenow fragment with DNA. Proc. Natl. Acad. Sci. USA 85 (1988) 8924–8928.

Gibbs, J.S., Chiou, H.C., Hall, J.D., Mount, D.W., Retondo, M.J., Weller, S.K. and Coen, D.M.: Sequence and mapping analysis of the herpes simplex virus DNA polymerase gene predict a C-terminal substrate binding domain. Proc. Natl. Acad. Sci. USA 82 (1985) 7969–7973.

Gibbs, J.S., Chiou, H.C., Bastow, K.F., Cheng, Y.-C. and Coen, D.M.: Identification of amino acids in herpes simplex virus DNA polymerase involved in substrate and drug recognition. Proc. Natl. Acad. Sci. USA 85 (1988) 6672–6676.

Hershfield, M.S.: On the role of deoxyribonucleic acid polymerase in determining mutation rates. Characterization of the defect in the T4 deoxyribonucleic acid polymerase caused by the tsL88 mutation. J. Biol. Chem. 248 (1973) 1417–1423.

Horiuchi, T., Maki, H. and Sekiguchi, M.: A new conditional lethal mutator (*dnaQ49*) in *Escherichia coli* K-12. Mol. Gen. Genet. 163 (1978) 277–283.

Huberman, J.A.: New views of the biochemistry of eucaryotic DNA replication revealed by aphidicolin, an unusual inhibitor of DNA polymerase α. Cell 23 (1981) 647–648.

Jacobsen, H., Klenow, H. and Overgaard-Hansen, K.: The N-terminal amino acid sequences of DNA polymerase I of *Escherichia coli* and of the large and the small fragments obtained by a limited proteolysis. Eur. J. Biochem. 45 (1974) 623–627.

Joyce, C.M. and Steitz, T.A.: DNA polymerase I: from crystal structure to function via genetics. Trends Biochem. Sci. 12 (1987) 288–292.

Joyce, C.M., Ollis, D.L., Rush, J., Steitz, T.A., Konigsberg, W.H. and Grindley, N.D.F.: Relating structure to function for DNA polymerase I of *Escherichia coli*. UCLA Symp. Mol. Cell. Biol. 32 (1985) 197–205.

Jung, G., Leavitt, M.C., Hsieh, J.-C. and Ito, J.: Bacteriophage PRD1 DNA polymerase: evolution of DNA polymerases. Proc. Natl. Acad. Sci. USA 84 (1987) 8287–8291.

Kempken, F., Meinhardt, F. and Esser, K.: In organello replication and viral affinity of linear, extrachromosomal DNA of the ascomycete *Ascobolus immersus*. Mol. Gen. Genet. 218 (1989) 523–530.

Khan, N.N., Wright, G.E., Dudycz, L.W. and Brown, N.C.: Butylphenyl dGTP: a selective and potent inhibitor of mammalian DNA polymerase alpha. Nucleic Acids Res. 12 (1984) 3695–3706.

Khan, N.N., Wright, G.E., Dudycz, L.W. and Brown, N.C.: Elucidation of the mechanism of selective inhibition of mammalian DNA polymerase alpha by 2-butylanilinopurines: development and characterization of 2-(*p*-*n*-butylanilino) adenine and its deoxyribonucleotides. Nucleic Acids Res. 13 (1985) 6331–6342.

Knopf, C.W.: The herpes simplex virus type 1 DNA polymerase gene: site of phosphonoacetic acid resistance mutation in strain Angelotti is highly conserved. J. Gen. Virol. 68 (1987) 1429–1433.

Kornberg, A: DNA Replication. Freeman, San Francisco, CA, 1980.

Larder, B.A., Kemp, S.D. and Darby, G.: Related functional domains in virus DNA polymerases. EMBO J. 6 (1987) 169–175.

Lawyer, F.C., Stoffel, S., Saiki, R.K., Myambo, K., Drummond, R. and Gelfand, D.H.: Isolation, characterization and expression in *Escherichia coli* of the DNA polymerase gene from *Thermus aquaticus*. J. Biol. Chem. 264 (1989) 6427–6437.

Leavitt, M.C. and Ito, J.: T5 DNA polymerase: structural-functional relationships to other DNA polymerases. Proc. Natl. Acad. Sci. USA 86 (1989) 4465–4469.

Lehman, I.R. and Chien, J.R.: Persistence of deoxyribonucleic acid polymerase I and its associated 5′ → 3′ exonuclease activity in *polA* mutants of *Escherichia coli* K12. J.Biol. Chem. 248 (1973) 7717–7713.

López, P., Martinez, S., Diaz, A., Espinosa, M. and Lacks, S.A.: Characterization of the polA gene of *Streptococcus pneumoniae* and comparison of the DNA polymerase I it encodes to homologous enzymes from *Escherichia coli* and phage T7. J. Biol. Chem. 264 (1989) 4255–4263.

Matsumoto, K., Takano, H., Kim, C.I. and Hirokawa, H.: Primary structure of bacteriophage M2 DNA polymerase: conserved segments within protein-priming DNA polymerases and DNA polymerase I of *Escherichia coli*. Gene 84 (1989) 247–255.

Morrison, A., Christensen, R.B., Alley, J., Beck, A.K., Bernstine, E.G., Lemontt, J.F. and Lawrence, C.W.: *REV3*, a *Saccharomyces cerevisiae* gene whose function is required for induced mutagenesis, is predicted to encode a non-essential DNA polymerase. J. Bacteriol. 171 (1989) 5659–5667.

Mullen, G.P., Shenbagamurthi, P. and Mildvan, A.S.: Substrate and DNA binding to a 50-residues peptide fragment of DNA polymerase I. J. Biol. Chem. 264 (1989) 19637–19647.

Oeser, B. and Tudzynski, P.: The linear mitochondrial plasmid pCIK1 of the phytopathogenic fungus *Claviceps purpurea* may code for a DNA polymerase and an RNA polymerase. Mol. Gen. Genet. 217 (1989) 132–140.

Ollis, D.L., Brick, R., Hamlin, R., Xuong, N.G. and Steitz, T.A.: Structure of the large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. Nature 313 (1985a) 762–766.

Ollis, D.L., Kline, C. and Steitz, T.A.: Domain of *E. coli* DNA polymerase I showing sequence homology to T7 DNA polymerase. Nature 313 (1985b) 818–819.

Pandey, V.N., Williams, K.R., Stone, K.L. and Modak, M.J.: Photoaffinity labeling of the thymidine triphosphate binding domain in *Escherichia coli* DNA polymerase I: identification of histidine-881 as the site of cross-linking. Biochemistry 26 (1987) 7744–7748.

Reha-Krantz, L.J.: Amino acid changes coded by bacteriophage T4 DNA polymerase mutator mutants. Relating structure to function. J. Mol. Biol. 202 (1988) 711–724.

Reha-Krantz, L.J.: Locations of amino acid substitutions in bacteriophage T4 tsL56 DNA polymerase predict an N-terminal exonuclease domain. J. Virol. 63 (1989) 4762–4766.

Rush, J. and Konigsberg, W.H.: Photoaffinity labeling of the Klenow fragment with 8-azido-dATP. J. Biol. Chem. 265 (1990) 4821–4827.

Spicer, E.K., Rush, J., Fung, C., Reha-Krantz, L.J., Karam, J.D. and Konigsberg, W.H.: Primary structure of T4 DNA polymerase: evolutionary relatedness of prokaryotic and eukaryotic DNA polymerases. J. Biol. Chem. 263 (1988) 7478–7486.

Tabor, S. and Richardson, C.C.: Selective inactivation of the exonuclease activity of bacteriophage T7 DNA polymerase by in vitro mutagenesis. J. Biol. Chem. 264 (1989) 6447–6458.

Tommasino, M., Ricci, S. and Galeotti, C.L.: Genome organization of the killer plasmid pGKL2 from *Kluyveromyces lactis*. Nucleic Acids Res. 16 (1988) 5863–5878.

Wong, S.W., Wahl, A.F., Yuan, P.-M., Arai, N., Pearson, B.E., Arai, K.-i., Korn, D., Hunkapiller, M.W. and Wang, T.S.-F.: Human DNA polymerase α gene expression is cell proliferation dependent and its primary structure is similar to both prokaryotic and eukaryotic replicative DNA polymerases. EMBO J. 7 (1988) 37–47.